7

Abt

memorandum

Environmental Research Area

4800 Montgomery Lane, Suite 600 ● Bethesda, MD 20814-5341 ● (301)-913-0500

Abt Associates Inc.

Date: January 13, 2003

To: Lynn Zipf, USEPA/Office of Water

From: Susan Keane, Brad Firlie, Kristina Watts and Emily Connor, Abt Associates

Subject: Procedure to Modify RSEI Model for Environmental Justice Indicators
Work Assignment 3-11 under Contract No. 68-C99-239

The Risk Screening Environmental Indicator (RSEI) model uses effluent data reported in the Toxic Release Inventory (TRI) to estimate the potential environmental impact of industrial releases. The current version of the model does not allow examination of differences in impacts among race groups or income groups resulting from releases to water bodies. A significant difference in impacts among race categories or among groups with different incomes may indicate a disparity in environmental justice. This memo briefly describes the component of the RSEI model related to water releases, and outlines the steps required to modify the water component in order to examine environmental justice issues.

How RSEI Models Exposures from Water Releases

The RSEI model analyzes the amount of a chemical released, the toxicity of the chemical, its fate and transport through the environment, the route and extent of human exposure, and the number of people affected, to create numerical values ("scores") that describe relative risk. The model divides the geographical area of the U.S. into 1 km by 1 km grid cells. Data from the U.S. Census is used to define the number of people living in each of these cells, and the demographic characteristics of these people.

For water releases, RSEI assumes that releases are discharged into the stream reach nearest the facility, unless the actual receiving body is known, in which case the actual reach is used. In each stream reach, chemical concentrations are modeled as the chemicals travel downstream up to 200 kilometers. The concentration in fish in these reaches is also estimated. Based on these estimates, two main pathways of human exposure are considered: fish ingestion and drinking water consumption.

Fish ingestion

The model uses the estimated fish concentrations and assumptions regarding fish consumption rates in recreational and subsistence fishing populations to estimate potential impacts on these fish consumers. The model also estimates the size and age/gender characteristics of the population of recreational and subsistence fishers potentially affected by a release. First, a county-level data set containing the number of fishing or hunting/fishing combination licenses was created from state fish and wildlife licensing data for 1996 (if 1996 was not available, 1997 was used). The number of fishing licenses in a county was then divided by the 1996 total population in the county. If no licensing information for a county was available, all of the grid cells in that county were assigned the ratio of total state licenses to total state population. If no information was available for the state in which the grid cell is located, the ratio for the state closest to that grid cell is assigned. The resulting ratio was then multiplied by the population in each grid cell to obtain the number of individuals with fishing licenses within each cell. To account for family members who also eat fish caught by one member, the model multiplies the number of fishers by 2.65, the size of the average U.S. household in 1995 (U.S. Census Bureau, 1996). Next, the population that consumes fish is then apportioned based on whether fish are eaten recreationally or for subsistence. Recreational fishers may fish during only certain times of the year for recreational purposes or to supplement their diet. In contrast, subsistence fishers may fish throughout the year and a major part of their diets may consist of fish they catch. Data are lacking on numbers of recreational compared to subsistence fishers. The model assumes that of the population that eats non-commercial fish, 95 percent eat fish on a recreational basis, and the remaining 5 percent subsist on fish.

The fishing population in each cell is then assigned to specific stream reaches where they are presumed to catch fish. This is done in two steps. First, every grid cell in the continental U.S. and its territories is analyzed to identify the stream reaches within 80 kilometers (50 miles) directly to the north, south, east, and west of the cell, equivalent to a 161 x 161 km square. This distance is based on a finding reported in the *1991 National Survey of Fishing, Hunting, and Wildlife-Associated Recreation* that 65 percent of anglers travel less than 50 miles to fish (U.S. Department of the Interior, 1993). Second, the fish-eating population in the grid cell is apportioned to each stream reach within 80 km surrounding the cell based on the ratio of the length of that reach to the total reach kilometers that exist within this distance. For example, Reach A and B may be located within 80 km of a given cell. If Reach A is 15 km in length and Reach B is 5 km in length (and the entire length of each reach is completely within 80 km of the cell), then a total of 20 km of stream reaches are located within the specified distance. Because Reach A represents three-fourths (15/20) and Reach B represents one-fourth (5/20) of total reach-kilometers within the specified range, the model therefore assumes that three-fourths of the fishing population in the cell catches fish from Reach A and one-fourth catches fish from Reach B. Note that the model uses only the portion of the reach's length that is inside the 161 by 161 km square area.

By apportioning the fishers to all nearby reaches, the method accounts for *all* possible fishing areas (i.e., stream reaches) within 80 km of the grid cell. The model then matches the chemical concentration in fish in the appropriate reach to the correctly-apportioned population. This is done for all reaches that have modeled chemical concentrations. Note that due to the complexity of this procedure, this procedure has currently only been applied to 1990 Census data; however, the same procedure can be applied to 2000 Census data as well.

Drinking water

To estimate the size of the population exposed to TRI releases through drinking water, the model uses estimates of the population served by each drinking water intake from the Safe Drinking Water Information System (SDWIS). However, this data set only lists the intake location and the number of people served by the water system. In many cases, there are multiple water intakes per water system. In the absence of other data, it is assumed that the total population of the water system is exposed to the full concentration of the released chemical estimated at each water intake.

The drinking water intake information from SDWIS contains only the number of people served by each drinking water system; it does not provide demographic or locational information for those served. To derive demographic information (age and sex breakdowns) for the population served, RSEI first identifies all of the people located in grid cells within 50 miles of each drinking water intake. The percentage of people in each of the ten age-sex categories is calculated for the entire area. Then, these percentages are applied to the SDWIS intake population, creating subpopulation groups that are used for calculating results.

Results of the RSEI model and Implications for Environmental Justice

The fundamental unit of analysis in the RSEI model is the "Indicator Element," which reflects the relative impact of a given release of a particular chemical from a particular reporting facility. These elements can be aggregated, so that users can look at all impacts from a given facility, from a set of facilities in a given industry or industries, from all facilities that are located in a given geographic area, or from all facilities releasing a particular chemical or chemicals. Regardless of the manner in which these results are aggregated, it should be emphasized that the underlying "Elements," from which aggregate results are compiled, are associated with a particular source of release. Specifically, RSEI can provide aggregate results by geographic area, but this aggregation reflects only the impacts of facilities *located in that area*; it does <u>not</u> consider the impacts on that area resulting from releases from facilities in other locations. For example, in the current model, the total risk-related score for all water releases in Indiana is attributable to water releases from that state's 485 TRI-reporting facilities. However, this score does not account for the impacts on population caused by chemicals flowing into the state in streams from other states. Furthermore, the RSEI score may include impacts that occur in other nearby states but are caused by facilities in Indiana.

In contrast, for EJ analysis, it would be preferred to select a geographic area of concern, and then sum the impacts of all of the releases affecting that area, regardless of their location of origin. The capacity currently exists to perform this kind of analysis for air releases; however, this capability has not been developed yet for water releases.

Also, as described above, the population data used by RSEI for drinking water pathway does not contain specific demographic information on populations served. RSEI makes very rough estimates of age and sex characteristics for these populations, but an accurate analysis of the exposure of specific populations through drinking water is not possible using the SDWIS data set. This limitation would also apply to estimates of race and income characteristics. Therefore, EJ analysis based on the consumption of fish would be the most feasible at this time. Additional possible approaches for future consideration of the drinking water pathway are described below.

Steps to Modify the Water Component for Environmental Justice Analysis

To develop the EJ component for water, the current RSEI model would require two major modifications: the program would have to be rewritten to calculate the cumulative impacts of all releases on a given geographic area; and the Census data set would be updated to include 2000 data, including race and income data. In addition, it may also be desirable to examine the relationship between race/income and behavior factors that affect exposure (i.e., subsistence fishing rate) and to incorporate consideration of such relationships, if they exist, into the model. Note that for now the modified model does not include drinking water exposure.

1. Modify model programming. Several significant changes to the model design must be made. First, the relevant release information must be extracted from the RSEI data set. The modified model will only include information related to releases to water (direct water, POTW transfer, wastewater treatment, and off-site wastewater treatment). Second, the execution of the model must be changed to save "cell-level" results. Currently, as the program is executed, it follows the release of a chemical from a facility downstream, and "accumulates" the impacts on each cell until all downstream impacts are accounted for. Only this summed result (which is attributed back to the responsible facility) is saved as the "Indicator Element" for that release. In order to look at the total impacts in one geographic location, the impacts calculated for each cell must be saved individually so they can be aggregated by cell, rather than by facility. Implementing a cell level aggregation requires changing the orientation of exposed individuals from the stream reach to each cell. Third, the capability to calculate the cumulative impacts on a given cell from all releases must be added to the program. Finally, an interface must be created that allows the user to query these results.

2. Modify Population Data to Include Race and Income. Currently, race or income level categories are not included in the population data used in RSEI (demographic categories only include age and sex). This data must be added to the model. Also, as described above, the apportionment of fishers to streams has only been conducted for 1990 data; this apportionment procedure needs to be applied to 2000 Census data as well.

The U.S. Census has relatively recently released 2000 data related to income and race. Income and race data are available at the block group level and block level, respectively. Income data are expressed in the following categories:

Less than $10,000
$10,000 to $14,999
$15,000 to $19,999
$20,000 to $24,999
$25,000 to $29,999
$30,000 to $34,999
$35,000 to $39,999
$40,000 to $44,999
$45,000 to $49,999
$50,000 to $59,999
$60,000 to $74,999
$75,000 to $99,999
$100,000 to $124,999
$125,000 to $149,999
$150,000 to $199,999
$200,000 or more

The racial and ethnic classifications used in the 2000 Census differ significantly than those used in the 1990 Census. The 2000 Census allowed respondents to select more than one category for race, listed questions on race in a new sequence, and added a new race category by dividing "Asian and Pacific Islander" into two categories. Respondents in the 2000 Census could choose one or more of the following six racial categories: (1) White, (2) Black or African American, (3) American Indian or Alaskan Native, (4) Asian, (5) Native Hawaiian or Other Pacific Islander, and (6) Some other race. Using these six basic race categories, 63 possible racial classifications exist, including six classifications for respondents who select only one race and 57 different classifications for respondents who select two or more races.

To analyze EJ issues, these race and income data would be incorporated into the population files of the RSEI model at the grid cell level. These data would become another "demographic characteristic" of the fishing populations that have already been assigned in RSEI. Note that resources may be required to purchase or license these data sets in a processed format, as the US Census only releases these data to the public in a very cumbersome form that is difficult to use.

Entering 2000 Census race and income data alone would be adequate for cross-sectional analysis; however, for trends analysis, we would have to enter such data for both the 1990 and 2000 Census data, and perform interpolation for intervening years. For trends analysis, one issue that must be resolved is how to compare the race categories between the 1990 and 2000 Census; in the 1990 Census, the population was categorized into only 5 racial categories, whereas the 2000 Census allowed for any combination of these categories (as described above). For consistency, these new categories must be aggregated into the 1990 format. During the conference call on January 6, EPA (Dr. Nick Bouwes) indicated that the academic community may have reached some consensus on the procedure for this conversion. We look forward to receiving additional information/references for this procedure.

3. Investigate the Possible Relationship between Exposure Pathway Data and Income/Race

Currently the RSEI model assumes that 5 percent of any fishing population will be subsistence fishers. However, as discussed during the conference call, low-income populations may have a higher rate of subsistence fishing. More information is needed to relate the percentage of a population dependent on subsistence fishing to income level. Abt Associates will investigate the fish consumption surveys that underlie the fish consumption rates to determine whether income levels can be related to amount of fish consumption. Race may also play an important role in the rate of subsistence fishing; Asian-American and Native Americans have been shown in several studies to have higher rates of freshwater fish consumption than other population groups; we may also consider how to incorporate these data into the model.

Types of Analysis Possible with Resulting Module

The new module will allow users to evaluate impacts on given geographic areas, and to evaluate the relative impacts on different income groups and racial categories. These results can be analyzed to evaluate disparate impacts on low-income groups or selected racial groups. We will also investigate the feasibility of retaining the ability to attribute the impacts in a particular geographic area to TRI release sources.

Options for Consideration of Additional Exposure through Drinking Water

As discussed during the conference call, and as described above, it is most feasible now to focus on EJ analysis through the fish consumption pathway. Unless data becomes available that identifies the location of the populations served by each water system, the model cannot easily assign impacts to individual cells with accuracy. However, there may be ways to make general assumptions regarding the location of populations served. For example, we could use the locations (latitude/longitude) of each treatment plant (if such data are available given current national security issues), and assume that populations in a given area around the treatment plant are served by that plant. The size of the area considered around each plant would vary by size of the population served, and would depend on the average population density in that area. We could then "assign" the drinking water impacts from that plant to all the cells included in the given area; these impacts could be "added" to the impacts of fish ingestion in those cell. However, we emphasize that the feasibility of this or alternative approaches for drinking water require greater consideration and discussion, and any approach chosen would certainly require more resources than the implementation of the fish ingestion pathway.